

McGRAW-HILL  
BOOK COMPANY  
New York  
St. Louis  
San Francisco  
Düsseldorf  
Johannesburg  
Kuala Lumpur  
London  
Mexico  
Montreal  
New Delhi  
Panama  
Rio de Janeiro  
Singapore  
Sydney  
Toronto

ALEXANDER M. MOOD  
*Professor of Administration and  
Director of Public Policy Research Organization  
University of California, Irvine*

FRANKLIN A. GRAYBILL  
*Department of Statistics  
Colorado State University  
Fort Collins, Colorado*

DUANE C. BOES  
*Department of Statistics  
Colorado State University  
Fort Collins, Colorado*

# Introduction to the Theory of Statistics

THIRD EDITION

EXHIBIT

A

Blumberg No. 5208

## cGRAW-HILL SERIES IN PROBABILITY AND STATISTICS

AVID BLACKWELL and HERBERT SOLOMON, *Consulting Editors*

BARUCHA-REID: Elements of the Theory of Markov Processes and Their Applications  
BAKE: Fundamentals of Applied Probability Theory  
BRENFELD AND LITTAUER: Introduction to Statistical Methods  
IBBONS: Nonparametric Statistical Inference  
RAYBILL: Introduction to Linear Statistical Models, Volume I  
EFFREY: The Logic of Decision  
I: Introduction to Experimental Statistics  
ILLER: Simultaneous Statistical Inference  
LOOD, GRAYBILL, AND BOES: Introduction to the Theory of Statistics  
MORRISON: Multivariate Statistical Methods  
EARCE: Biological Statistics: An Introduction  
EIFFER: Concepts of Probability Theory  
LAJ: The Design of Sample Surveys  
LAJ: Sampling Theory  
THOMASIAN: The Structure of Probability Theory with Applications  
WADSWORTH AND BRYAN: Application of Probability and Random Variables  
WASAN: Parametric Estimation  
WEISS: Statistical Decision Theory  
WOLF: Elements of Probability and Statistics

CONTENTS

To HARRIET A.M.M.  
To my GRANDCHILDREN F.A.G.  
To JOAN, LISA, and KARIN D.C.B.

Library of Congress Cataloging in Publication Data

Mood, Alexander McFarlane, 1913-  
Introduction to the theory of statistics.  
(McGraw-Hill series in probability and statistics)  
Bibliography: p.  
I. Mathematical statistics. I. Graybill,  
Franklin A., joint author. II. Boes, Duane C.,  
joint author. III. Title.  
QA276.M67 1974 519.5 73-292  
ISBN 0-07-042864-6

INTRODUCTION  
TO THE THEORY  
OF STATISTICS

Copyright © 1963, 1974 by McGraw-Hill, Inc. All rights reserved.  
Copyright 1950 by McGraw-Hill, Inc. All rights reserved.  
Printed in the United States of America. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means, electronic,  
mechanical, photocopying, recording, or otherwise, without the prior written  
permission of the publisher.

121314 KPKP 8987654

This book was set in Times Roman.  
The editors were Brete C. Harrison and Madelaine Eichberg;  
the cover was designed by Nicholas Krenitsky;  
and the production supervisor was Ted Agrillo.  
The drawings were done by Oxford Illustrators Limited.  
The printer and binder was Kinsport Press, Inc.

Preface to the Third Edition	xi
Excerpts from the First and Second Edition Prefaces	x
I Probability	
1 Introduction and Summary	
2 Kinds of Probability	
2.1 Introduction	
2.2 Classical or a Priori Probability	
2.3 A Posteriori or Frequency Probability	
3 Probability—Axiomatic	
3.1 Probability Models	
3.2 An Aside—Set Theory	
3.3 Definitions of Sample Space and Event	
3.4 Definition of Probability	
3.5 Finite Sample Spaces	
3.6 Conditional Probability and Independence	

test with critical region  $= \{(x_1, \dots, x_n): \sum (x_i - \mu)^2 > k^*\}$  is uniformly most powerful of size  $\alpha$ , where  $k^*$  is given by  $P_{\sigma^2 = \sigma_0^2}[\sum (X_i - \mu)^2 > k^*] = \alpha$ , which implies that  $k^* = \sigma_0^2 \chi_{1-\alpha}^2(n)$ , where  $\chi_{1-\alpha}^2(n)$  is the  $(1 - \alpha)$ th quantile point of the chi-square distribution with  $n$  degrees of freedom.

If  $\mu$  is unknown, a test can be found using the statistic  $V = \sum (X_i - \bar{X})^2 / \sigma_0^2$ .  $V$  will tend to be larger for  $\sigma^2 > \sigma_0^2$  than for  $\sigma^2 \leq \sigma_0^2$ ; so a reasonable test would be to reject  $\mathcal{H}_0$  for  $V$  large. If  $\sigma^2 = \sigma_0^2$ , then  $V$  has a chi-square distribution with  $n - 1$  degrees of freedom, and  $P_{\sigma^2 = \sigma_0^2}[V > \chi_{1-\alpha}^2(n - 1)] = \alpha$ , where  $\chi_{1-\alpha}^2(n - 1)$  is the  $(1 - \alpha)$ th quantile of a chi-square distribution with  $n - 1$  degrees of freedom. It can be shown that the test given by the following: Reject  $\mathcal{H}_0$  if and only if  $\sum (X_i - \bar{X})^2 / \sigma_0^2 > \chi_{1-\alpha}^2(n - 1)$  is a generalized likelihood-ratio test of size  $\alpha$ .

$\mathcal{H}_0: \sigma^2 = \sigma_0^2$  versus  $\mathcal{H}_1: \sigma^2 \neq \sigma_0^2$  We leave the case  $\mu$  assumed known as an exercise. For  $\mu$  unknown, so that  $\Theta_0 = \{(\mu, \sigma): -\infty < \mu < \infty; \sigma^2 = \sigma_0^2\}$ , we can find a size- $\alpha$  test using the confidence-interval method. In Subsec. 3.2 of Chap. VIII, we found the following  $100\gamma$  percent confidence interval for  $\sigma^2$ :

$$\left( \frac{(n-1)S^2}{q_2}, \frac{(n-1)S^2}{q_1} \right),$$

where  $q_1$  and  $q_2$  are quantile points of a chi-square distribution with  $n - 1$  degrees of freedom, say  $f_Q(q; n - 1)$ , satisfying

$$\int_{q_1}^{q_2} f_Q(q; n - 1) dq = \gamma.$$

A size- $(\alpha = 1 - \gamma)$  test is given by the following: Accept  $\mathcal{H}_0$  if and only if  $\sigma_0^2$  is contained in the above confidence interval. It is left as an exercise to show that for a particular pair of  $q_1$  and  $q_2$  the test of size  $\alpha$  derived by the confidence-interval method is in fact the generalized likelihood-ratio test of size  $\alpha$ .

### 4.3 Tests on Several Means

In this subsection we will consider testing hypotheses regarding the means of two or more normal populations. We begin with a test of the equality of two means.

**Equality of two means** In many situations it is necessary to compare two means when neither is known. If, for example, one wished to compare two proposed new processes for manufacturing light bulbs, one would have to base the comparison on estimates of both process means. In comparing the yield of

a new line of h  
estimates of bc  
the standard l  
would be grow  
the same seaso  
mean yields fo  
is thus speciali  
over a period

The gene  
a random vari  
a random vari  
two samples, c

$$\mathcal{H}_0: \mu_1 = \mu_2,$$

The parameter  
 $X_2$  is specified  
The subspace  
 $(\mu, \sigma_1^2, \sigma_2^2)$  nec  
under the hyp

We shal  
the sample fr  
from the secon

$$L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$$

$$= \left( \frac{1}{2\pi\sigma_1^2} \right)^n$$

and its maxim

$$\sup_{\underline{\theta}} L = \left[ - \right. \\ \left. 2\tau \right]$$

If we put  $\mu_1$  an  
and  $\sigma_2^2$ , it wil  
equation and  
generalized lik  
find its distrib  
variances. TI

uniformly most  
 $k^*] = \alpha$ , which  
 tile point of the

$\sum (X_i - \bar{X})^2 / \sigma_0^2$ .  
 able test would  
 are distribution  
 $1)] = \alpha$ , where  
 ion with  $n - 1$   
 the following:  
 neralized likeli-

ed known as an  
 $\sigma^2 = \sigma_0^2$ , we  
 Subsec. 3.2 of  
 rval for  $\sigma^2$ :

$$\frac{S^2}{\sigma^2},$$

th  $n = 1$  degrees

$= \gamma$ .

und only if  $\sigma_0^2$  is  
 ise to show that  
 the confidence-  
 size  $\alpha$ .

ie means of two  
 y of two means.

o compare two  
 o compare two  
 ld have to base  
 ring the yield of

a new line of hybrid corn with that of a standard line, one would also have to use estimates of both mean yields because it is impossible to state the mean yield of the standard line for the given weather conditions under which the new line would be grown. It is necessary to compare the two lines by planting them in the same season and on the same soil type and thereby obtain estimates of the mean yields for both lines under similar conditions. Of course the comparison is thus specialized; a complete comparison of the two lines would require tests over a period of years on a variety of soil types.

The general problem is this: We have two normal populations—one with a random variable  $X_1$ , which has a mean  $\mu_1$  and variance  $\sigma_1^2$ , and the other with a random variable  $X_2$ , which has a mean  $\mu_2$  and variance  $\sigma_2^2$ . On the basis of two samples, one from each population, we wish to test the null hypothesis

$$\mathcal{H}_0: \mu_1 = \mu_2, \sigma_1^2 > 0, \sigma_2^2 > 0 \quad \text{versus} \quad \mathcal{H}_1: \mu_1 \neq \mu_2, \sigma_1^2 > 0, \sigma_2^2 > 0.$$

The parameter space  $\bar{\Theta}$  here is four-dimensional; a joint distribution of  $X_1$  and  $X_2$  is specified when values are assigned to the four quantities  $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ . The subspace  $\bar{\Theta}_0$  is three-dimensional because values for only three quantities  $(\mu, \sigma_1^2, \sigma_2^2)$  need be specified in order to specify completely the joint distribution under the hypothesis that  $\mu_1 = \mu_2 = \mu$ , say.

We shall suppose that there are  $n_1$  observations  $(X_{11}, X_{12}, \dots, X_{1n_1})$  in the sample from the first population and  $n_2$  observations  $(X_{21}, X_{22}, \dots, X_{2n_2})$  from the second. The likelihood function is

$$L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}) = L \\
= \left( \frac{1}{2\pi\sigma_1^2} \right)^{n_1/2} \exp \left[ -\frac{1}{2} \sum_1^{n_1} \left( \frac{x_{1i} - \mu_1}{\sigma_1} \right)^2 \right] \left( \frac{1}{2\pi\sigma_2^2} \right)^{n_2/2} \exp \left[ -\frac{1}{2} \sum_1^{n_2} \left( \frac{x_{2j} - \mu_2}{\sigma_2} \right)^2 \right],$$

and its maximum in  $\bar{\Theta}$  is readily seen to be

$$\sup_{\bar{\Theta}} L = \left[ \frac{n_1}{2\pi \sum_1^{n_1} (x_{1i} - \bar{x}_1)^2} \right]^{n_1/2} \left[ \frac{n_2}{2\pi \sum_1^{n_2} (x_{2j} - \bar{x}_2)^2} \right]^{n_2/2} e^{-n_1/2} e^{-n_2/2}.$$

If we put  $\mu_1$  and  $\mu_2$  equal to  $\mu$ , say, and try to maximize  $L$  with respect to  $\mu$ ,  $\sigma_1^2$ , and  $\sigma_2^2$ , it will be found that the estimate of  $\mu$  is given as the root of a cubic equation and will be a very complex function of the observations. The resulting generalized likelihood-ratio  $\lambda$  will therefore be a complicated function, and to find its distribution is a tedious task indeed and involves the ratio of the two variances. This makes it impossible to determine a critical region  $0 < \lambda < k$

for a given probability of a Type I error because the ratio of the population variances is assumed unknown. A number of special devices can be employed in an attempt to circumvent this difficulty, but we shall not pursue the problem further here. For large samples the following criterion may be used: The root of the cubic equation can be computed in any instance by numerical methods, and  $\lambda$  can then be calculated; furthermore, as we shall see in Sec. 5 below, the quantity  $-2 \log \Lambda$  has approximately the chi-square distribution with one degree of freedom, and hence a test that would reject for  $-2 \log \lambda$  large could be devised.

When it can be assumed that the two populations have the same variance, the problem becomes relatively simple. The parameter space  $\bar{\Theta}$  is then three-dimensional with coordinates  $(\mu_1, \mu_2, \sigma^2)$ , while  $\bar{\Theta}_0$  for the null hypothesis  $\mu_1 = \mu_2 = \mu$  is two-dimensional with coordinates  $(\mu, \sigma^2)$ . In  $\bar{\Theta}$  we find that the maximum-likelihood estimates of  $\mu_1, \mu_2$ , and  $\sigma^2$  are, respectively,  $\bar{x}_1, \bar{x}_2$ , and

$$\frac{1}{n_1 + n_2} \left[ \sum_{i=1}^{n_1} (x_{1i} - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2 \right];$$

so

$$\sup_{\bar{\Theta}} L = \left\{ \frac{n_1 + n_2}{2\pi \left[ \sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2 \right]} \right\}^{(n_1 + n_2)/2} e^{-(n_1 + n_2)/2}.$$

In  $\bar{\Theta}_0$ , the maximum-likelihood estimates of  $\mu$  and  $\sigma^2$  are

$$\hat{\mu} = \frac{1}{n_1 + n_2} \left( \sum_{i=1}^{n_1} x_{1i} + \sum_{j=1}^{n_2} x_{2j} \right) = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} \quad \text{for } \mu$$

and

$$\begin{aligned} & \frac{1}{n_1 + n_2} \left[ \sum (x_{1i} - \hat{\mu})^2 + \sum (x_{2j} - \hat{\mu})^2 \right] \\ &= \frac{1}{n_1 + n_2} \left[ \sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2 + \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)^2 \right] \\ & \quad \text{for } \sigma^2, \end{aligned}$$

which gives

$$\begin{aligned} & \sup_{\bar{\Theta}_0} L \\ &= \left[ \frac{n_1 + n_2}{2\pi \left[ \sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2 + \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)^2 \right]} \right]^{(n_1 + n_2)/2} \\ & \quad \times e^{-(n_1 + n_2)/2}. \end{aligned}$$

Fin:

Sub  
qua  
entl  
 $\sigma^2/\mu$   
 $\mu_1 - \bar{X}_1$   
inde  
deg  
with

is n

has  
hav  
aliz

and  
cou  
 $T^2$   
t di

$\mathcal{H}_1$   
 $T >$   
(1 -

Equ  
just  
ava

Finally,

$$\lambda = \left( 1 + \frac{[n_1 n_2 / (n_1 + n_2)] (\bar{x}_1 - \bar{x}_2)^2}{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2} \right)^{-(n_1 + n_2)/2}. \quad (17)$$

This last expression is very similar to the corresponding one obtained in Subsec. 4.1, and it turns out that this test can also be performed in terms of a quantity which has the  $t$  distribution. We know that  $\bar{X}_1$  and  $\bar{X}_2$  are independently normally distributed with means  $\mu_1$  and  $\mu_2$  and with variances  $\sigma^2/n_1$  and  $\sigma^2/n_2$ . Also it is readily seen that  $\bar{X}_1 - \bar{X}_2$  is normally distributed with mean  $\mu_1 - \mu_2$  and variance  $\sigma^2(1/n_1 + 1/n_2)$ . Under the null hypothesis the mean of  $\bar{X}_1 - \bar{X}_2$  will be 0. The quantities  $\sum (X_{1i} - \bar{X}_1)^2/\sigma^2$  and  $\sum (X_{2j} - \bar{X}_2)^2/\sigma^2$  are independently distributed as chi-square distributions with  $n_1 - 1$  and  $n_2 - 1$  degrees of freedom, respectively; hence their sum has the chi-square distribution with  $n_1 + n_2 - 2$  degrees of freedom. Since under the null hypothesis

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{1/n_1 + 1/n_2}}$$

is normally distributed with mean 0 and unit variance, the quantity

$$T = \frac{\sqrt{n_1 n_2 / (n_1 + n_2)} (\bar{X}_1 - \bar{X}_2)}{\sqrt{[\sum (X_{1i} - \bar{X}_1)^2 + \sum (X_{2j} - \bar{X}_2)^2] / (n_1 + n_2 - 2)}} \quad (18)$$

has the  $t$  distribution with  $n_1 + n_2 - 2$  degrees of freedom. [Note that we do have independence of the numerator and denominator in Eq. (18).] The generalized likelihood-ratio is

$$\lambda = \left[ \frac{1}{1 + [t^2 / (n_1 + n_2 - 2)]} \right]^{(n_1 + n_2)/2}, \quad (19)$$

and its distribution is determined by the  $t$  distribution. The test would, of course, be done in terms of  $T$  rather than  $\lambda$ . A 5 percent critical region for  $T$  is  $T^2 > [t_{.975}(n_1 + n_2 - 2)]^2$ , where  $t_{.975}(n_1 + n_2 - 2)$  is the .975th quantile of the  $t$  distribution with  $n_1 + n_2 - 2$  degrees of freedom.

If we want to test  $\mathcal{H}_0: \mu_1 = \mu_2$  versus  $\mathcal{H}_1: \mu_1 > \mu_2$  or  $\mathcal{H}_0: \mu_1 \leq \mu_2$  versus  $\mathcal{H}_1: \mu_1 > \mu_2$ , a size- $\alpha$  test is given by the following: Reject  $\mathcal{H}_0$  if and only if  $T > t_{1-\alpha}(n_1 + n_2 - 2)$ , where  $T$  is defined in Eq. (18) and  $t_{1-\alpha}(n_1 + n_2 - 2)$  is the  $(1 - \alpha)$ th quantile of the  $t$  distribution with  $n_1 + n_2 - 2$  degrees of freedom.

**Equality of several means** The test presented above can be extended from just two normal populations to  $k$  normal populations. We assume that we have available  $k$  random samples, one from each of  $k$  normal populations; that is,